

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-098219

(43)Date of publication of application : 09.04.1999

(51)Int.Cl.

H04L 29/08

H04L 12/56

H04L 13/08

(21)Application number : 09-259474

(71)Applicant : NIPPON TELEGR & TELEPH CORP <NTT>
SONY CORP

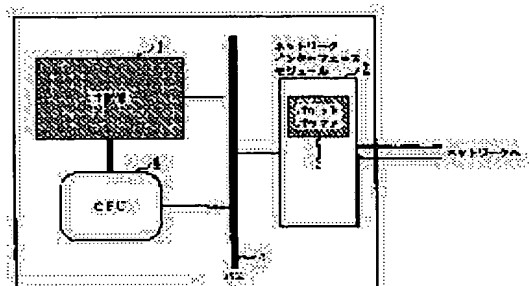
(22)Date of filing : 25.09.1997

(72)Inventor : YAMASHITA TAKASHI
TANAKA HIROYUKI
UCHIUMI HIDESUKE
SHIONOZAKI ATSUSHI(54) COMMUNICATION METHOD HAVING COMMUNICATION PROTOCOL FOR OPERATING INTER-LAYER
FLOW CONTROL AND DATA COMMUNICATION TERMINAL

(57)Abstract:

PROBLEM TO BE SOLVED: To prevent packet losses from being generated in a terminal and to prevent throughput or network use efficiency from deteriorating.

SOLUTION: A data communication protocol for operating flow control between a transport layer and a data link layer is realized. (1) When the activity ratio of a packet buffer 3 in a data link layer is over a first threshold value, the data link layer communicates the interrupt of packet transfer to a transport layer. When the activity ratio is not higher than a second threshold value and smaller than the first threshold value, the data link layer requests the resumption of packet transfer to the transport layer. (2) At the time of operating packet transfer from the transport layer to the lower rank layer, packet transfer is operated when the packet amount is not more than a value decided from the transfer speed and the maximum burst size, and when the the packet amount is more than the value, the packet transfer is operated after the lapse of a decided time, and the speed is suppressed to a constant value.



LEGAL STATUS

[Date of request for examination]

17.05.2001

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the
examiner's decision of rejection or application
converted registration]

[Date of final disposal for application]

[Patent number]

3413788

[Date of registration]

04.04.2003

[Number of appeal against examiner's decision of
rejection]

[Date of requesting appeal against examiner's decision]

This Page Blank (uspto)

of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

This Page Blank (uspto)

【特許請求の範囲】

【請求項1】 上層から下層に向けてアプリケーション層、プレゼンテーション層、セッション層、トランスポート層、ネットワーク層、データリンク層および物理層の階層構造を持ち、データ通信プロトコルに従う通信方法において、上記トランスポート層からネットワーク層を通してデータリンク層へパケットを転送する際に、該データリンク層のパケットバッファでバッファ溢れが生じ易い場合には、該データリンク層がパケット転送を一時中止させるように該トランスポート層に通知するか、あるいは、

該データリンク層のパケット送出速度を該トランスポート層に通知し、該トランスポート層が通知された速度以下の速度でパケットを下位層に転送することにより、該トランスポート層とデータリンク層との間でフロー制御を行う通信プロトコルを持つことを特徴とする通信方法。

【請求項2】 請求項1に記載の通信方法において、前記フロー制御では、データリンク層内のパケットバッファの使用率が第1のしきい値を上回ったとき、該データリンク層がトランスポート層にパケット転送を中止するよう通知し、

その後、上記パケットバッファの使用率が上記第1のしきい値より小さい第2のしきい値を下回ったとき、該データリンク層が該トランスポート層にパケット転送の再開を要求することを特徴とする通信方法。

【請求項3】 請求項1に記載の通信方法において、前記フロー制御では、トランスポート層で時刻 t にパケットを下位層に転送しようとしたとき、時刻 $t-s$ から t までに送出されたパケットの総データ量 P_{sum} と、送出しようとしているパケットのサイズ P との和 $P_{sum}+P$ が最初のパケット転送の時刻以降の時刻 $t-s$ に対して、パケット転送速度 R_t と最大バーストサイズ B_{max} から決められる値 $R_t \times (t-s) + B_{max}$ 以下になる場合にのみパケットを転送し、それ以外の場合には、上記条件を満たすも以降の時刻 t' まで待ってからパケット転送を行うことにより、トランスポート層からデータリンク層へのパケット転送速度を、データリンク層のパケット転送速度以下の一定値 R_t に抑えることを特徴とする通信方法。

【請求項4】 中央演算器と主記憶装置にバス接続されたネットワークインターフェースモジュールを持つデータ通信端末において、請求項1～請求項3のいずれかに記載のフロー制御を上記中央演算器上で実行することを特徴とするデータ通信端末。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、トランスポート層とデータリンク層との間でパケットを受け渡すデータ通

信プロトコルを持つ通信方法およびそのデータ通信プロトコルを用いて通信を行うデータ通信端末に関する。

【0002】

【従来の技術】 TCP/IP (Transmission Control Protocol/Internet Protocol) や OSI (Open Systems Interconnection 開放型システム間相互接続) プロトコルのような、データ通信用のインターフェース条件であるデータ通信プロトコルは、図2に示すような階層構造を有している。すなわち、アプリケーションプログラムが実行する動作レベルで、端末の特性や符号化制御等を内容とするアプリケーション層16、通信応用プロファイルや文書応用プロファイル等の内容を持つプレゼンテーション層15、発着ID等の制御機能を持つセッション層14、上位層のための網に依存しないトランスポート手順を行うトランスポート層13、各ネットワークの呼設定、切断等の制御を行うネットワーク層12、データリンクの確立、解放、エラー制御等を行うデータリンク層11、およびユーザーと網電圧物理インタフェースの条件を設定する物理層10を有している。パケットを送出する場合、アプリケーション層16（以下では、アプリケーション層16に、プレゼンテーション層15やセッション層14も含めて考えることにする）のソフトウェアで生成されたパケットは、トランスポート層13を実現するソフトウェアに渡された後、さらにトランスポート層13からネットワーク層12へ、ネットワーク層12からデータリンク層11へと、下位層へ順次渡されていく。最終的に、パケットは電気や光のパルスのような形態で表現されることにより、端末の外部に送出される。

【0003】 この場合、通常の実装では、パケットのメモリコピーが実際に行われるのは、2回だけである。すなわち、図1に示すように、主記憶1内でアプリケーション層16からトランスポート層13へパケットが移動される時と、データリンク層11内で端末の主記憶1からバス5を介してネットワークインターフェースモジュール2内のパケットバッファ3へパケットが移動される時である。データリンク層11より上のその他の層の間でのパケットの移動は、パケットが収納されている主記憶1内の位置へのポインターを受け渡すだけで実現される。アプリケーション層16からトランスポート層13へのメモリコピーの場合には、フロー制御が行われるため、トランスポート層13内のパケットバッファ溢れによるデータ損失は生じない。すなわち、トランスポート層13内のパケットバッファが溢れそうになると、アプリケーション層16がパケットのコピーを一時中止するという制御が行われる。しかし、データリンク層11と上位層との間では、フロー制御は行われない。データリンク層11が持つパケットバッファには、ネットワークインターフェースモジュール2内の他に、主記憶1内の

ものもあるが、フロー制御が行われないため、両方のバッファが溢れても上位層がパケットをデータリンク層11に転送し続け、パケット廃棄が生じてしまう。

【0004】上述のような状態が生じている理由は、従来のデータ通信プロトコルの基本設計方針として、トランスポート層13より下の層ではデータ通信の信頼性を保証せず、また各層の間の相互作用が最小限になるようにしていたためである。データ通信の信頼性は、トランスポート層13でのパケット再送のみにより保証されていた。従って、データリンク層11内のパケットバッファに空きが無い場合でも、トランスポート層13のプロトコルはパケット送出を中止せずにパケットを損失させ、後からパケットを再送していた。しかし、従来においては、アプリケーション層16でパケットを生成して、トランスポート層13やネットワーク層12を通してデータリンク層11へ転送する速度が、ネットワークインターフェースモジュール2が端末の外にパケットを送出する速度よりも遅かったため、データリンク層11内のパケットバッファ溢れによるパケット損失は殆んど生じなかった。

【0005】

【発明が解決しようとする課題】このように、従来のデータ通信プロトコルでは、端末の高性能化やネットワークの高機能化に伴って、データ通信のスループットが大きく劣化する可能性がある。すなわち、近年、中央演算装置等の端末の構成部品の性能が急激に上昇したため、上位層からデータリンク層11へパケットを転送する速度の方が、ネットワークインターフェースモジュール2が端末の外にパケットを送出する速度より速い場合が生じてきた。このような場合には、データリンク層11内のパケットバッファが溢れ、端末内であるにもかかわらず、パケットの損失が生じる可能性がある。また、データリンク層11が高機能化し、ATMのようにパケット送信速度を自由に変更できるものが用いられてきた。この場合、ネットワークの輻輳状況によりデータリンク層11から端末の外へのパケット送信速度を小さく抑えることもあり得る。そのようなとき、上位層からデータリンク層11へパケットを転送する速度との差がさらに大きくなるため、パケット損失の可能性が大きくなる。パケットの損失が生じると、トランスポート層13はパケット再送を行う必要があるため、データ送信スループットが低下し、またネットワークの帯域も浪費してしまう。特にTCPのような輻輳制御を行うトランスポートプロトコルの場合には、パケット損失を検出すると、送信ウィンドウサイズを小さくする等により自律的にパケット送信速度を低く抑えるため、著しいスループットの低下が生じる。

【0006】図3は、最大送信ウィンドウサイズを変化させた場合のTCPのスループットを示す特性図である。送信端末と受信端末の間の往復伝搬遅延(Round

trip Time, RTT)は、20msである。このRTTが大きい場合、TCPで高速なスループットを実現するためには、送信ウィンドウサイズを大きくする必要がある。送信ウィンドウサイズとは、着信確認応答を受けずに(一括)送信可能な最大データ長のことである。この送信ウィンドウサイズを大きくし過ぎると、データリンク層11内のパケットバッファが溢れてパケットの損失が生じる。これは、上位層が送信ウィンドウサイズ分のデータを一度にまとめてデータリンク層11へ転送しようとするためである。パケット損失が生じると、TCPの輻輳制御が働いて、スループットが大きく低下する。このように、従来のデータ通信プロトコルのままでは、本来避けられるべき端末内部でのパケット損失により、データ送信スループットの低下やネットワーク使用効率の低下が起きるという問題があった。そこで、本発明の目的は、このような従来の課題を解決し、トランスポート層とデータリンク層との間でパケットを転送する際のパケットの廃棄等をなくして、端末内でパケット損失を生じさせず、スループットやネットワーク使用効率の低下を避けることが可能な通信方法およびデータ通信端末を提供することにある。

【0007】

【課題を解決するための手段】上記目的を達成するため、本発明の通信方法では、トランスポート層とデータリンク層との間でフロー制御を行う。すなわち、データリンク層でバッファ溢れが生じないように、トランスポート層からのパケット転送を制御するのである。第1のフロー制御の方法は、データリンク層のパケットバッファの使用率が高くてバッファ溢れが生じ易い場合に、データリンク層がパケット転送を一時中止させるようにトランスポート層に通知する方法である。また、第2のフロー制御方法は、データリンク層のパケット送出速度をトランスポート層に通知し、トランスポート層が通知された速度以下の速度でパケットを下位層に転送する方法である。本発明によれば、トランスポート層とデータリンク層との間でフロー制御を行うため、端末内でパケット損失が生じることなく、スループットやネットワーク使用効率の低下を回避することができる。

【0008】

【発明の実施の形態】以下、本発明の実施例を、図面により詳細に説明する。図1は、本発明が適用されるデータ通信端末の構成図であり、図3は、フロー制御を行った場合のスループットを示す特性図である。図1に示すように、通信プロトコルはソフトウェアで実現され、中央演算装置4により実行される。送信されるパケットは、アプリケーション層により主記憶1上に生成される。次に、データリンク層を実現するソフトウェアにより、ネットワークインターフェースモジュール2内のパケットバッファ3に移動される。ネットワークインターフェースモジュール2によりパケットは端末の外に送出

される。図3の特性図では、本発明によるフロー制御つきTCPと従来のTCPとが比較のために図示されている。TCPのウィンドウ制御では、受信側のバッファの大きさに応じて連続して受信できるパケット数（ウィンドウサイズ）により送受信が制御される。従来のTCPでは、最大ウィンドウサイズが約170KB以上になると、スループットが低下している。それに対して、本発明によるフロー制御つきTCPでは、最大ウィンドウサイズが170KB以上になってもスループットは低下することなく上昇する。フロー制御つきTCPでは、トランスポート層からのデータ転送速度を、指定した速度に保つ機能を通常のTCPに付加した。指定した速度は、グラフに示された実際のスループットとほぼ同じである。通常のTCPとは異なり、送信ウィンドウサイズが大きくても、スループットが低下しないことが分る。これは、パケット損失が生じないためである。

【0009】（第1の実施例）図4は、第1の実施例を示すフロー制御方法の動作フローチャートである。第1の実施例では、データリンク層のパケットバッファの使用率が高くてバッファ溢れが生じ易い場合、データリンク層がパケット転送を一時中止させるようにトランスポート層に通知する方法（第1のフロー制御方法）を用いる。第1のフロー制御方法では、データリンク層を実現するソフトウェア（データリンク層プログラムと呼ぶ）からも、またトランスポート層を実現するソフトウェア（トランスポート層プログラムと呼ぶ）からもアクセスできる、バッファ溢れの可能性を示す変数（これをバッファ溢れフラグと呼ぶ）を主記憶1内に用意する。図4（a）はデータリンク層プログラムの動作を示し、図4（b）はトランスポート層プログラムの動作を示している。データリンク層内のパケットバッファの使用率がしきい値 T_1 を上回った場合に、データリンク層プログラムはバッファ溢れフラグを立てる（ステップ201）。一方、トランスポート層プログラムは、パケットを下位レイヤに転送する前にバッファ溢れフラグを検査し（ステップ211）、フラグが立っているか否かを判別して、フラグが立っていたならばパケット送出を中止する（ステップ212）。その後、データリンク層内のパケットバッファの使用率がしきい値 T_2 を下回った時点で、データリンク層プログラムはバッファ溢れフラグをおろし（ステップ203、204）、トランスポート層プログラムを呼び出してパケット送出の再開を依頼する（ステップ205）。ただし、しきい値 T_1 、 T_2 は、ネットワークインターフェースモジュールでの特性によって定まる値（ $T_1 \geq T_2$ ）である。トランスポート層プログラムは、データリンク層からパケット送出再開の依頼が来たことを知ると（ステップ214）、パケット送出を再開する（ステップ215）。そして、送出すべきパケットがなくなるまで（ステップ216）、データリンク層のバッファ溢れフラグが立っていないことを確

認しながらパケットを送出する（ステップ211、213）。

【0010】（第2の実施例）第2の実施例では、データリンク層のパケット送出速度をトランスポート層に通知し、トランスポート層が通知された速度以下の速度でパケットを下位層に転送する方法（第2のフロー制御方法）を用いる。このフロー制御方法では、まず、データリンク層プログラムは、データ送信開始時に、トランスポート層プログラムにデータリンク層のデータ送信速度を通知する。また、データ送信開始後も、データリンク層のデータ送信速度が変化する度にデータリンク層のデータ送信速度を通知する。トランスポート層プログラムは、その内部にトラヒックシェーピング機能、つまり下位層へのデータ転送速度を制御する機能を持ち、データリンク層のデータ送信速度以下でパケット転送を行う。すなわち、トランスポート層で時刻 t にパケットを下位層に転送しようとしたとき、時刻 $t-s$ から t までに送出されたパケットの総データ量 P_{sum} と、送出しようとしているパケットのサイズ P との和 $P_{sum}+P$ が最初のパケット転送の時刻以降の時刻 $t-s$ に対して、パケット転送速度 R_t と最大バーストサイズ B_{max} から決められる値 $R_t \times (t-s) + B_{max}$ 以下になる場合にのみパケットを転送し、それ以外の場合には、上記条件を満たす t 以降の時刻 t' まで待ってからパケット転送を行うことにより、トランスポート層からデータリンク層へのパケット転送速度を、データリンク層のパケット転送速度以下の一定値 R_t に抑える。

【0011】図5は、本発明の第2の実施例を示すトラヒックシェーピング機能のフローチャートである。図5の実施例では、上述の第2のフロー制御方法に関して、送信判断時の判定式中のパケット送信開始時刻 $t-s$ から現在時刻 t までの間の任意の時刻 s_0 の範囲を限定して実施した場合を考えている。トランスポート層プログラムは、データ転送速度 R_t 、最大バーストサイズ B_{max} の2つのパラメータを持ち、トークンという変数 T を持っている。トークンの初期値は最大バーストサイズ B_{max} に等しい。概念的には、トークンはデータ転送速度 R_t の割合で時間とともに増加するが、最大バーストサイズ B_{max} 以上にはならないものとする。アイドル状態にした後（ステップ100）、アプリケーション層からパケットが到着すると（ステップ101）、先ずトークン T を更新する（ステップ103）。すなわち、 $T = \min(B_{max}, T + R_t \times (t - t_p))$ の値に更新する。ここで、 t は現在の時刻であり、 t_p は前回のパケット送出時刻である。次に、パケットサイズ P とトークン T とを比較し（ステップ104）、 $P < T$ であればパケットを送出する（ステップ106）。パケットを送出すると、トークン T はパケットサイズ P だけ減少される（ $T = T - P$ ）。パケット送出した後、送信すべきパケットがあるか否かを判別し（ステップ10

7)、なければアイドル状態に戻り(ステップ100)、あればタイムアウトにする(ステップ102)。
 【0012】一方、TとPの比較の結果、Pに比べてトークンTが足りない場合にはパケットは送出されず、トークンTがP以上になる時刻を計算して($t_0 = t + (P - T) / R_t$)、その時刻までのタイムアウトを設定する(ステップ105)。タイムアウトが起きると、タイムアウト処理に入り(ステップ102)、トランスポート層プログラムが呼び出され、パケット転送が再開される(ステップ101)。なお、データ転送速度 R_t は、データリンク層から通知されるデータリンク層のデータ送信速度 R_1 以下でなければならない。また、データリンク層でのバッファ溢れを避けるため、最大バーストサイズ B_{max} は、 $Buf \times (R_{max} - R_1) / R_{max}$ 以下にする必要がある。ここで、 Buf は、データリンク層内のパケットバッファの大きさ、 R_{max} はトランスポート層から下位層への最大転送速度である。
 【0013】
 【発明の効果】以上説明したように、本発明によれば、トランスポート層とデータリンク層との間でフロー制御を行うため、端末内でパケット損失が生じることがな

く、スループットの低下とネットワーク使用効率の低下を避けることができる。

【図面の簡単な説明】

【図1】本発明が適用されるデータ通信端末の構成図である。

【図2】データ通信プロトコルの階層構造を示す図である。

【図3】本発明によるスループットと従来のTCPによるスループットを比較した特性図である。

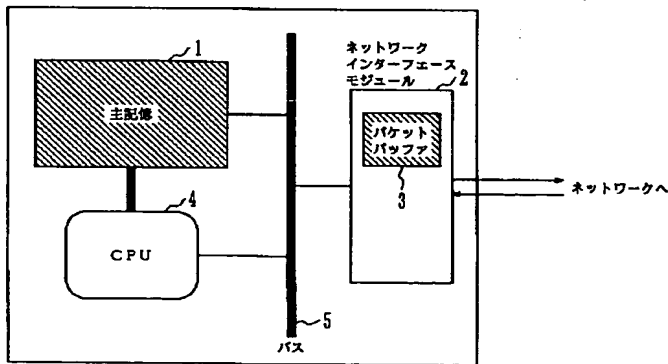
【図4】本発明の第1の実施例を示すフロー制御方法の動作フローチャートである。

【図5】本発明の第2の実施例を示すトラヒックシェーピング機能のフローチャートである。

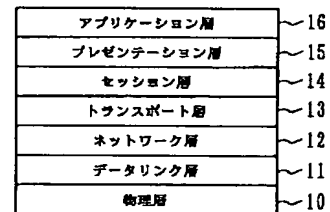
【符号の説明】

1…主記憶、2…ネットワークインターフェースモジュール、3…パケットバッファ、4…中央演算装置(CPU)、5…バス、10…物理層、11…データリンク層、12…ネットワーク層、13…トランスポート層、14…セッション層、15…プレゼンテーション層、16…アプリケーション層。

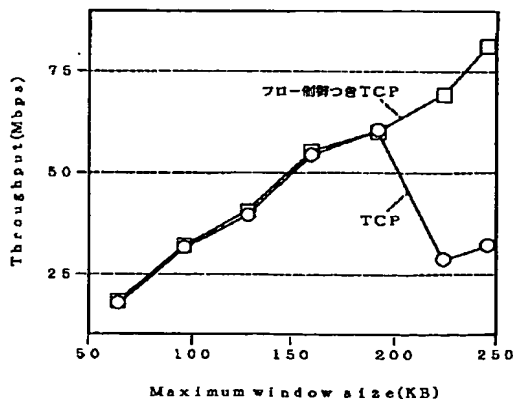
【図1】



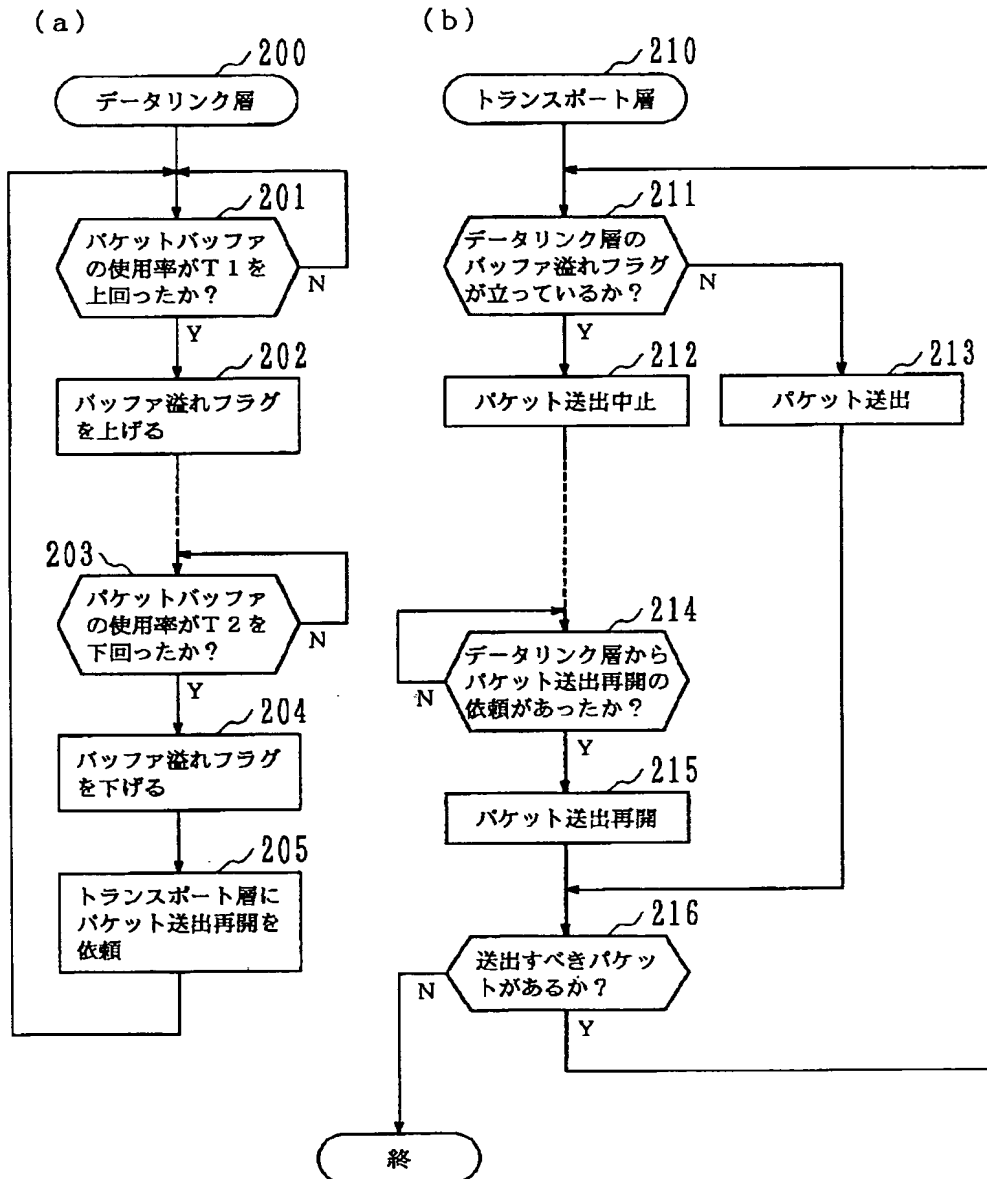
【図2】



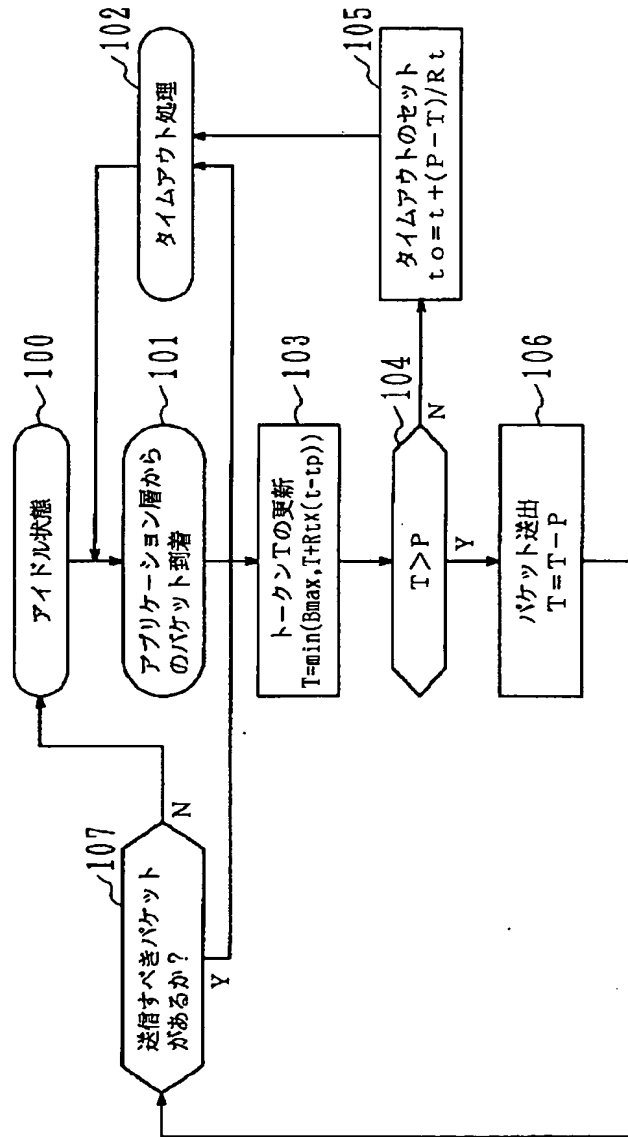
【図3】



【図4】



【図5】



R_t : トランスポート層のデータ転送速度
 B_{max} : 最大バーストサイズ
 T : トークン
 P : パケットサイズ
 t : 現在の時刻
 t_p : 前回のパケット送出時刻
 t_o : タイムアウトする時刻
 $R_t \leq R_1$
 $B_{max} \leq Buf \times (R_{max} - R_1) / R_{max}$
 R_1 : リンク層のデータ送信速度
 Buf : リンク層内のパケットバッファのサイズ

フロントページの続き

(72)発明者 内海 秀介
 東京都品川区北品川6丁目7番35号 ソニ
 ー株式会社内

(72)発明者 塩野崎 敦
 東京都品川区北品川6丁目7番35号 ソニ
 ー株式会社内

This Page Blank (uspto)